In Press at Journal of Experimental Psychology: General

The effects of repetition on belief in naturalistic settings

Lisa K. Fazio, Raunak M. Pillai, Deep Patel

Vanderbilt University

Author Note

Lisa K. Fazio, Department of Psychology and Human Development, Vanderbilt University. Raunak M. Pillai, Department of Psychology and Human Development, Vanderbilt University. Deep Patel, Department of Psychology and Human Development, Vanderbilt University.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1937963. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This research was also supported by a grant from Facebook Research. The authors would like to thank Sean Polyn for his assistance with the computational modeling. Any errors are ours alone. All data, materials, modeling code and additional analyses are available online at the project's OSF site, along with our preregistration of the analyses and sample size: https://osf.io/re6dh.

Correspondence concerning this article should be addressed to Lisa K. Fazio, 230 Appleton Place #552, Jesup 105, Nashville, TN 37203. E-mail: lisa.fazio@vanderbilt.edu. Phone: 615-875-8301. Fax: 615-343-9494.

Abstract

In our modern well-connected world, false information spreads quickly and is often repeated multiple times. From laboratory studies, we know that this repetition can be harmful as repetition increases belief. However, it is unclear how repetition affects belief in real-world settings. Here we examine a larger number of repetitions (16), more realistic timing of the repetitions (across two weeks), and more naturalistic exposures (text messages). 435 US participants recruited from mTurk were texted true and false trivia statements across 15 days before rating the accuracy of each statement. Statements were seen either 1, 2, 4, 8, or 16 times. We find clear evidence that repetition increases belief. Initial repetitions produced the largest increase in perceived truth, but belief continued to increase with additional repetitions. We introduce a simple computational model suggesting that current accounts are insufficient to explain this observed pattern and that additional theoretical assumptions (e.g., that initial repetitions are more strongly encoded) are required. Practically, the results imply that repeated exposure to false information during daily life can increase belief in that misinformation.
Keywords: truth, repetition, illusory truth, misinformation

The effects of repetition on belief in naturalistic settings

Scams, fakes, misinformation and disinformation have a long history of causing a variety of monetary and societal harms. However, in our current well-connected world, misinformation can travel faster and further than ever before. Social media platforms help millions of people communicate and stay connected, but they also allow false information to spread rapidly (e.g., Vosoughi et al., 2018). In the recent US election, a Georgia poll worker went into hiding because of viral false claims that he destroyed an absentee ballot (Brumback & Joffe-Block, 2020) and election officials were overwhelmed with phone calls from people concerned about false voting-related conspiracy theories and scenarios (Browning & Alba, 2020). This spread is troubling since simply hearing a statement multiple times increases people's belief in its accuracy. Called the illusory truth effect, this phenomenon has been studied since the 1970s and it helps explain why politicians, advertisers and propagandists often repeat their statements (see Dechêne et al., 2010; Unkelbach et al., 2019 for reviews).

Despite decades of research, we still know very little about how repetition affects belief in real-world situations. The typical lab-based paradigm involves comparing belief in statements encountered twice to those encountered once in a short 15 - 30-minute experimental session (see Fazio, 2020; Schwartz, 1982; Unkelbach & Greifeneder, 2018 for typical examples). However, what we actually care about is how repetition in news broadcasts, advertisements and social media affects belief in daily life. Repetition in the real world occurs in larger numbers and over longer timescales. On social media, people rarely see the same piece of misinformation twice in a row; instead, it is spread out across multiple days or weeks. In addition, misinformation is often seen many times, rather than just once or twice. For example, President Trump has repeated some of his most popular false claims hundreds of times (Washington Post Fact Checker, 2020).

In the following experiment, we use a novel texting paradigm to explore how repetition affects belief in naturalistic settings.

Examining the effect of larger numbers of repetitions is not only practically important. Our design also provides additional information and constraints on theoretical explanations of the illusory truth effect. With only two measurement points it is unclear if the effect of repetition on belief is linear, logarithmic or quadradic. It is only with additional measurements that we can accurately identify the shape of the relation between repetition and perceived truth and possible underlying cognitive mechanisms.

Prior research

A few researchers have examined the effects of multiple repetitions on belief rather than simply comparing novel and repeated statements. Most of these studies have only examined the influence of up to three exposures (Gigerenzer, 1984; Hasher et al., 1977; Pennycook et al., 2018), but a few studies have examined five (Hawkins et al., 2001), six (Arkes et al., 1991) or ten (DiFonzo et al., 2016) exposures[1]. Some studies fail to find an increase in truth ratings after a second presentation (Arkes et al., 1991), while others have found that belief continues to increase logarithmically up to nine repetitions (DiFonzo et al., 2016). In the current study, we will examine the effects of up to 16 exposures to expand our knowledge of how belief changes with multiple repetitions.

One difference between the previous experiments is the timing of the repetitions. In the studies that have found a continuing increase in belief (DiFonzo et al., 2016; Hawkins et al., 2001), the repetitions occurred within minutes. When the repetition was more spaced out (once a

[1] See also Hassan & Barber (2021) which was published as this manuscript was in the peer review process.

week) belief quickly plateaued (Arkes et al., 1991). A single repetition affects belief whether the exposures are separated by a few minutes (Fazio, 2020) or as long as 3 months (Brown & Nix, 1996). However, it's less clear if the size of the effect differs across time. An older meta-analysis found no differences between the size of the illusory truth effect across different delays (Dechêne et al., 2010), however a more recent registered report directly manipulated the delay and found that the effect was largest when repetitions occurred within minutes and smaller when the second exposure occurred days or weeks later (Henderson et al., 2021). To our knowledge, no prior research has directly examined how the timing of multiple repetitions may affect belief. Here, we distribute the repetitions over the span of two weeks to better match how real-life misinformation is experienced. In addition, we compare the effect of spaced (e.g., every four days) and massed (every day) repetition on belief.

Finally, existing research has not carefully examined the effects of repetition in real-world settings. Exposure to true and false trivia statements on posters hung around a college campus increased belief in those statements (Boehm, 1994), however only two statements were repeated and it is unclear how many times participants may have seen each poster. In contrast, we examine how repetition affects belief when participants receive information in text messages, a form of transmission much more similar to what occurs in daily life and which allows us to determine exactly how often each statement was viewed.

Present research

As described above, we extend current research on the illusory truth effect to examine the effects of multiple repetitions (up to 16) spread over a longer time period (two weeks, with exposures either spaced or on consecutive days) in a more real-world setting (exposure via test messages). Practically, this approach allows us to empirically test whether the illusory truth effect generalizes to more naturalistic settings and timelines of exposure. Relatedly, this work allows us to determine whether recent findings of a logarithmic increase in perceived truth will replicate with longer delays. In determining the shape of the curve between repetition and perceived truth, we are also able to explore whether existing explanations for how repetition increases perceived truth can be generalized to explain the effects of multiple repetitions.

To examine this latter question, we introduce a simple computational model formalizing theoretical assumptions in the illusory truth effect literature. Existing theories, based largely on experiments involving up to two or three exposures, identify two key assumptions: a) that repetition increases the familiarity and/or ease of processing of a statement (Unkelbach et al., 2019) and b) that people interpret greater than expected levels of these experiences as a cue for the truth of a statement (Hansen et al., 2008). Accordingly, we base our model on a prominent computational model of human memory and the experience of familiarity, MINERVA 2 (Hintzman, 1984). Critically, we test whether this model, and thereby the prominent assumptions stated in the literature, is sufficient to explain our data involving multiple repetitions, or whether additional assumptions are necessary. While our models are not full explanations of the illusory truth effect, they present a useful push towards generating more specific, concrete theories of how repetition affects perceived truth.

Method

Participants

Participants were recruited in two rounds from the Amazon mTurk Platform. Using the Cloud Research service, we restricted participation to users in the United States and blocked duplicate IP addresses (Litman et al., 2017). Participants agreed to receive 5 text messages per day for 15 days through the Remind platform (designed to help teachers text their students) and to complete a final survey. All participants provided informed consent and the study was conducted with the approval of the Vanderbilt Institutional Review Board (IRB # 190948). To incentivize participation, participants received $1 for initially signing up for Remind, $.05 for each text they responded to (75 texts = $3.75), $3 for completing the final survey and a $2.25 bonus if they replied to at least 95% of the texts. Thus, each participant could receive up to $10.

Our goal was to have approximately 500 participants in the final dataset. We predicted that one-third of the participants would not complete the full study, so we initially aimed to recruit 760 participants across the two rounds. For round one (May 2nd-16th 2020), only 182 participants of the 380 recruited followed though and signed up for the Remind platform. Thus, for round two (May 16th – 30th, 2020), we recruited 575 participants rather than the planned 380. 394 participants from this round signed up for the Remind platform, resulting in a total of 576 participants who received the text messages. Of those, 509 responded to at least half of the texts and were invited to the final survey (a preregistered requirement) and 436 completed the final survey. The final number of participants gives us 92% power, 95% CI [90%, 93%], to detect an effect of the log number of views on truth ratings of at least 0.05 with an alpha of .05 (computed using simr, Green & MacLeod, 2016).

Design

The experiment had a 2 (truth: true, false) x 5 (views: 1, 2, 4, 8, 16) within-subjects design. Half of the presented statements were true, and half were false. The number of repetitions for each statement varied with some statements appearing only on the final test and others being viewed 2, 4, 8, or 16 times across the study. Figure 1 shows an overview of the schedule for the key statements. To counterbalance which statement appeared in each schedule, we assigned participants to one of 10 counterbalancing groups, rotating the statement that appeared in each schedule across these groups. For the statements that were presented 2, 4, or 8 times, one statement was presented at equally spaced intervals across the 16 days and the other was presented on consecutive days immediately prior to the final test.

	Day															**Final survey**	
Statement	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	**Times viewed**
A	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	16
B	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	16
C		x		x		x		x		x		x		x		x	8
D									x	x	x	x	x	x	x	x	8
E				x				x				x				x	4
F													x	x	x	x	4
G								x								x	2
H															x	x	2
I																x	1
J																x	1

Figure 1. Presentation schedule for the key statements. Statements C, E, and G have spaced presentations, while statements D, F, and H have massed presentations. The spacing manipulation cannot be applied to statements A, B, I, and J.

Materials

Matching prior studies of the illusory truth effect (Fazio, 2020; Fazio et al., 2019), we used true and false trivia statements as the materials (Table 1). Twenty-eight statements were chosen from the unknown statements used by Fazio (2020). Other studies have shown that prior knowledge does not protect against the effects of repetition (Fazio, 2020; Fazio et al., 2015), so the results would likely generalize to other trivia statements.

Of the 28 statements, 10 were key statements (5 true, 5 false), 8 were novel fillers on the final test and were seen only in the final survey (4 true, 4 false), and 10 (5 true, 5 false) were filler items for days with less than 5 key statements scheduled for presentation. Only the 10 key statements are analyzed below.

Table 1.

Sample statements

True	False
In the story of Pinocchio, the goldfish is named Cleo.	Bach is the composer who wrote the opera "Don Giovanni."
Kingston is the capital of Jamaica.	Lima is the capital of Chile.
Bullet was the name of Roy Roger's dog.	Bell is the inventor of the wireless radio.

Procedure

After giving their informed consent, participants were randomly assigned to one of the 10 counterbalancing groups and received detailed instructions on how to sign up to receive the text messages. Participants were also informed that they would be receiving texts containing true and false trivia statements. Texts were sent at 11am, 1pm, 3pm, 5pm, and 7pm CDT daily. Each text contained a statement and asked participants how interested they were in the statement on a scale of 1 (low) to 6 (high), Figure 2. The messages were scheduled according to the presentation schedule seen in Figure 1. If five key statements were not being sent out that day, the remaining statements were fillers. For each day, the five statements were randomly assigned to the different time slots. Participants had 24 hours from the time our messages were sent to reply with their rating.

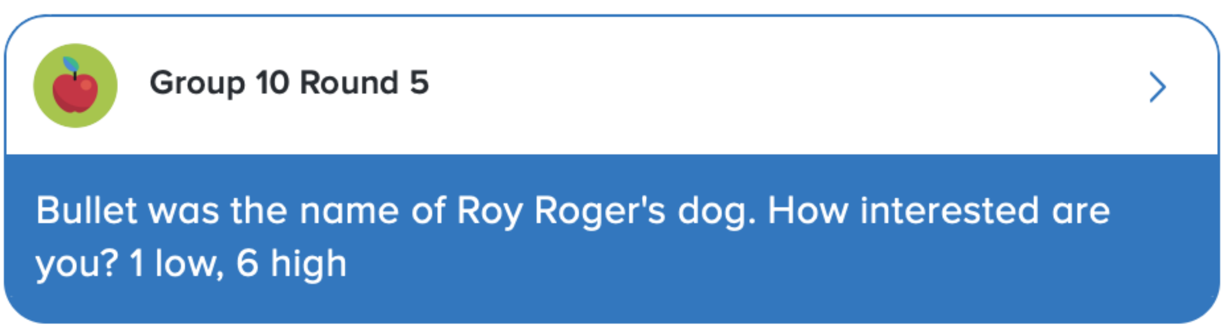

Figure 2. Sample text message with the rating scale.

On day 16, participants received an email link to a final survey that was available for 48 hours. This survey included all 10 key statements and eight additional novel filler statements (four true, four false).

Participants were asked to judge the accuracy of each statement and were told that "some will be true, while others will be not true." They were also asked not to google the statements, but to "just give whatever rating you think is correct." Participants were warned that some of the trivia facts would be ones they had seen before, while others would be new. For each statement, participants rated its truth on a 6-point scale (1 = Definitely False, 2 = Probably False, 3 = Possibly False, 4 = Possibly True, 5 = Probably True, 6 = Definitely True).

At the end of the survey, participants were asked whether they looked up any of the trivia statements online either while receiving the texts or answering the final survey (separate questions). If a participant answered "yes" to either question they were asked to list those statements. Finally, we asked participants what they thought the study was about.

Results

All data and additional analyses are available online at the project's OSF site, along with our preregistration of the analyses and sample size: https://osf.io/re6dh. All statistical analyses were conducted using the lme4 package (Bates et al., 2015) in R (R Core Team, 2020). Following our pre-registration, we excluded 48 truth ratings for statements that participants indicated looking up at any point during the experiment (1 - 10 statements across 25 participants, M = 1.92), leaving us with 4,312 truth ratings from 435 participants.[2]

Effects of the number of repetitions

[2] We also conducted exploratory analyses using only the high-response-rate participants. For these analyses, we included only the 273 participants who responded to at least 95% of the texts. The pattern of results is identical to the analyses with the full dataset reported here and can be viewed at the OSF repository https://osf.io/re6dh.

Our primary question was if repetition would affect belief in this naturalistic setting. As shown in Figure 3A, it did. Mean truth ratings increased with the number of repetitions and this increase appeared to be logarithmic. In fact, when the data was graphed as a function of the natural logarithm of the number of views, truth ratings demonstrated a linear increase with repetition (Figure 3B).

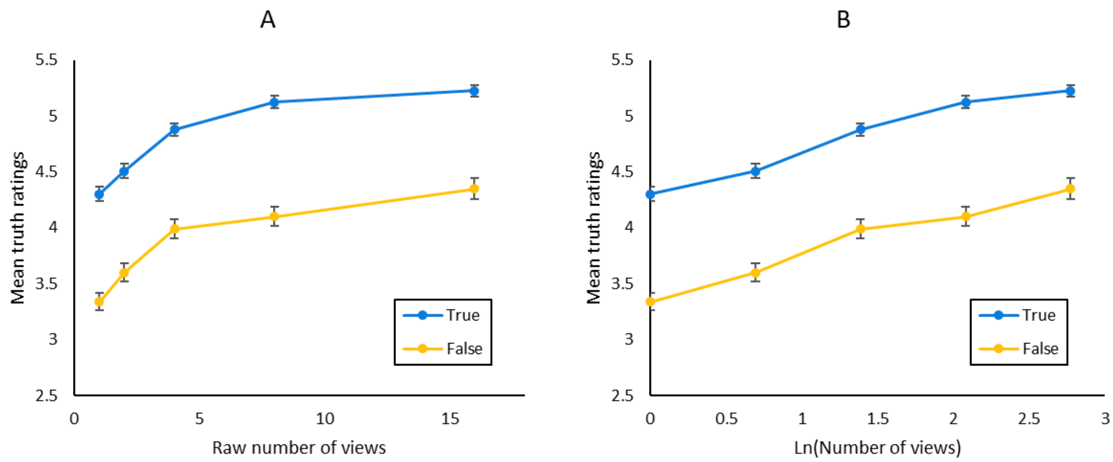

Figure 3. Mean truth ratings for true and false statements as a function of (A) the raw number of times viewed and (B) the natural logarithm of the number of views. Error bars reflect standard error of the mean. Participants responded on a scale of 1 = definitely false to 6 = definitely true.

As pre-registered, we evaluated these data statistically by fitting two mixed-effects linear regression models and comparing the model fit. In the first model, we used the raw number of views as a predictor variable, while the second model used the natural logarithm of the number of views. Using a likelihood-ratio test, the model using the natural logarithm of the number of views provided a better fit, so we describe only that model below.[3]

The final model predicted truth ratings as a function of the fixed effects of statement truth (true or false), the natural logarithm of the number of views (approximately: 0, 0.69, 1.39, 2.08,

[3] The results of the model using the raw number of views and the model comparison can be found in the OSF repository https://osf.io/re6dh.

or 2.77), and their interaction. Statement truth was contrast coded and centered at zero (true = -0.5 and false = 0.5), as was the log number of views (0 = -2, 0.69 = -1, 1.39 = 0, 2.08 = 1, and 2.77 = 2). An initial model with a maximal random effects structure failed to converge. Thus, as preregistered, we adopted a model which included by-subject and by-item estimates of the intercept only, Table 2.

Table 2.

Mixed-effects model with fixed effects of statement truth, natural logarithm of the number of views, and their interaction, as well as random intercepts for subjects and items.

Fixed Effects	Estimate	*SE*	*df*	*t* Value	*p* Value
Intercept	4.341	0.124	8.664	34.989	< .001
Ln(Repetition)	**0.251**	**0.015**	**3,871**	**16.869**	**< .001**
Truth	**-0.933**	**0.243**	**8**	**-3.837**	**.005**
Ln(Repetition)*Truth	0.006	0.030	3,870	0.228	.820
Random Effects	Variance	*SD*			
Subject (Intercept)	0.262	0.512			
Statement (Intercept)	0.144	0.379			

Note. Model was fit to 4,312 truth ratings from 435 participants across 10 statements. Bolded values indicate significant effects.

As shown in Table 2, we found a significant fixed effect of truth such that true statements (M = 4.81) were estimated to be rated 0.933 points higher than false statements (M = 3.87) on the 6-point truth scale (1 = definitely false, 6 = definitely true). We also found a significant fixed effect of repetitions such that an increase in one natural logarithmic unit of the number of views was estimated to correspond with a 0.251 increase in the statement's truth rating. Lastly, we found no interaction effect between statement truth and number of repetitions; repetition increased belief similarly for true and false statements.

Massed versus spaced repetition

Our next question was whether the spacing of the repetitions influenced participants' truth ratings. Because statements presented once or 16 times had only possible schedule, we restricted this analysis to the statements presented 2, 4, or 8 times. Given the clear superiority of the logarithmic model above, we modeled the effects using the natural logarithm of the number of views.

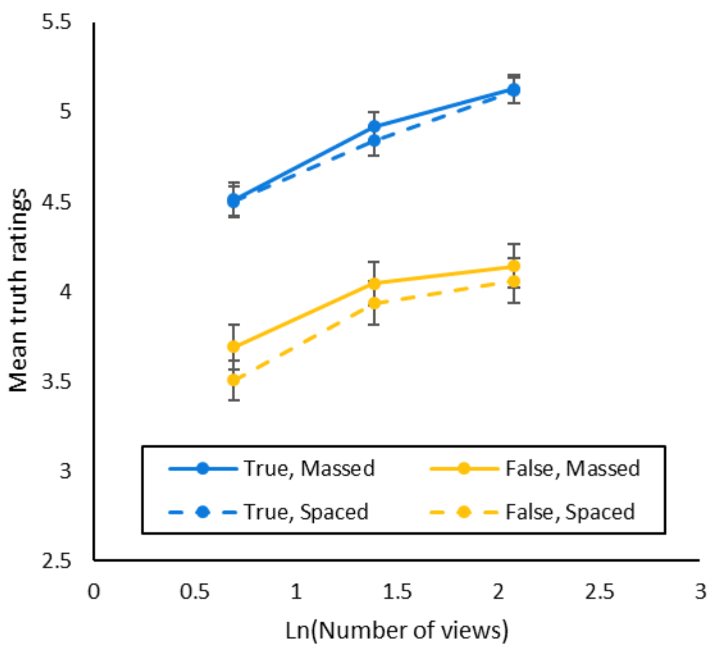

Figure 4. Mean truth ratings for true and false, massed and spaced statements as a function of the natural logarithm of the number of views. Error bars reflect standard error of the mean. Participants responded on a scale of 1 = definitely false to 6 = definitely true.

As shown in Figure 4, participants rated massed and spaced statements similarly; there was no evidence that the increase in truth ratings with repetition differed across the massed and spaced schedules. Following our preregistration, we added spacing (massed = -0.5, spaced = 0.5) as a predictor variable to the mixed-effect linear regression model above. We again initially fit a model with a maximal random effects structure, but, due to convergence errors, adopted a model which included by-subject and by-item estimates of the intercept only. As shown in Table 3, we found no significant effect of massed vs. spaced presentation schedules nor any significant interaction effects (all $p > .05$).

Table 3.

Mixed-effects model with fixed effects of statement truth, natural logarithm of the number of views, repetition schedule and all interactions, as well as random intercepts for subjects and items.

Fixed Effects	Estimate	*SE*	*df*	*t* Value	*p* Value
Intercept	4.365	0.112	8.671	37.71	< .001
Ln(Repetition)	**0.281**	**0.034**	**2147**	**8.301**	**< .001**
Truth	**-0.944**	**0.227**	**7.957**	**-4.168**	**.003**
Schedule	-0.079	0.055	2149	-1.423	.155
Ln(Repetition)*Truth	-.074	0.068	2148	-1.085	.278
Ln(Repetition)*Schedule	0.024	0.068	2148	0.356	.722
Truth*Schedule	-0.090	0.146	432.1	-0.618	.537
Ln(Repetition)*Truth*Schedule	0.054	0.135	2146	0.396	.692
Random Effects	Variance	*SD*			
Subject (Intercept)	0.245	0.495			
Statement (Intercept)	0.121	0.347			

Note. Model was fit to 2,591 truth ratings from 435 participants across 10 statements. Bolded values indicate significant effects.

Computational Modeling

We have demonstrated that repetition exerts a logarithmic increase on perceived truth—initial repetitions increase perceived truth to a greater extent than later repetitions. The key question is whether current theories can account for this pattern, or whether additional assumptions may be needed. To address this question, we introduce a computational model of the illusory truth effect. One benefit of computational modeling over verbal theories is that modeling forces you to be specific about your assumptions and assumed mechanisms. To be clear, we do not view the models below as complete models of the mechanisms behind the illusory truth effect, but we do think that the models are a useful step toward formalizing our theories and identifying possible mechanisms.

To this end, we developed and compared two computational models: Model 1, which minimally reflects current assumptions for explaining the illusory truth effect, and Model 2, which adds an extra assumption that successive repetitions of items are encoded into memory more poorly, likely because of a lack of attention. This latter assumption of novelty-sensitive encoding has been alluded to in previous research (cf. DiFonzo et al., 2016, p. 23), but has not been directly stated or tested. For example, the referential theory proposed by Unkelbach and Rom (2017) is the most recent (and one of the most detailed) theory of the illusory truth effect. According to the theory, people judge truth based on the cohesion between references in memory. With repetition, the references within a statement become more strongly connected, leading to increased truth judgments. However, the theory does not specify the specific relation between repetition and strengthened connections. It is possible that each repetition increases cohesion equally or that initial repetitions have a larger impact.

Further, standard illusory truth experiments in which items are only encoded once during an exposure phase cannot reveal the dynamics involved in multiple encodings of a single statement. Thus, it is essential to examine this assumption using data like those from the current experiment which include large numbers of repetitions. If existing theoretical explanations for the illusory truth effect (based largely on experiments with smaller numbers of repetitions) are sufficient to explain the effects of multiple repetitions, Model 1 should provide a sufficient fit to the data. Alternatively, an additional assumption, such as the novelty-sensitive encoding mechanism posited in Model 2, may be necessary to best explain the data. We discuss the models in detail below, and they are schematically depicted in Figure 5.

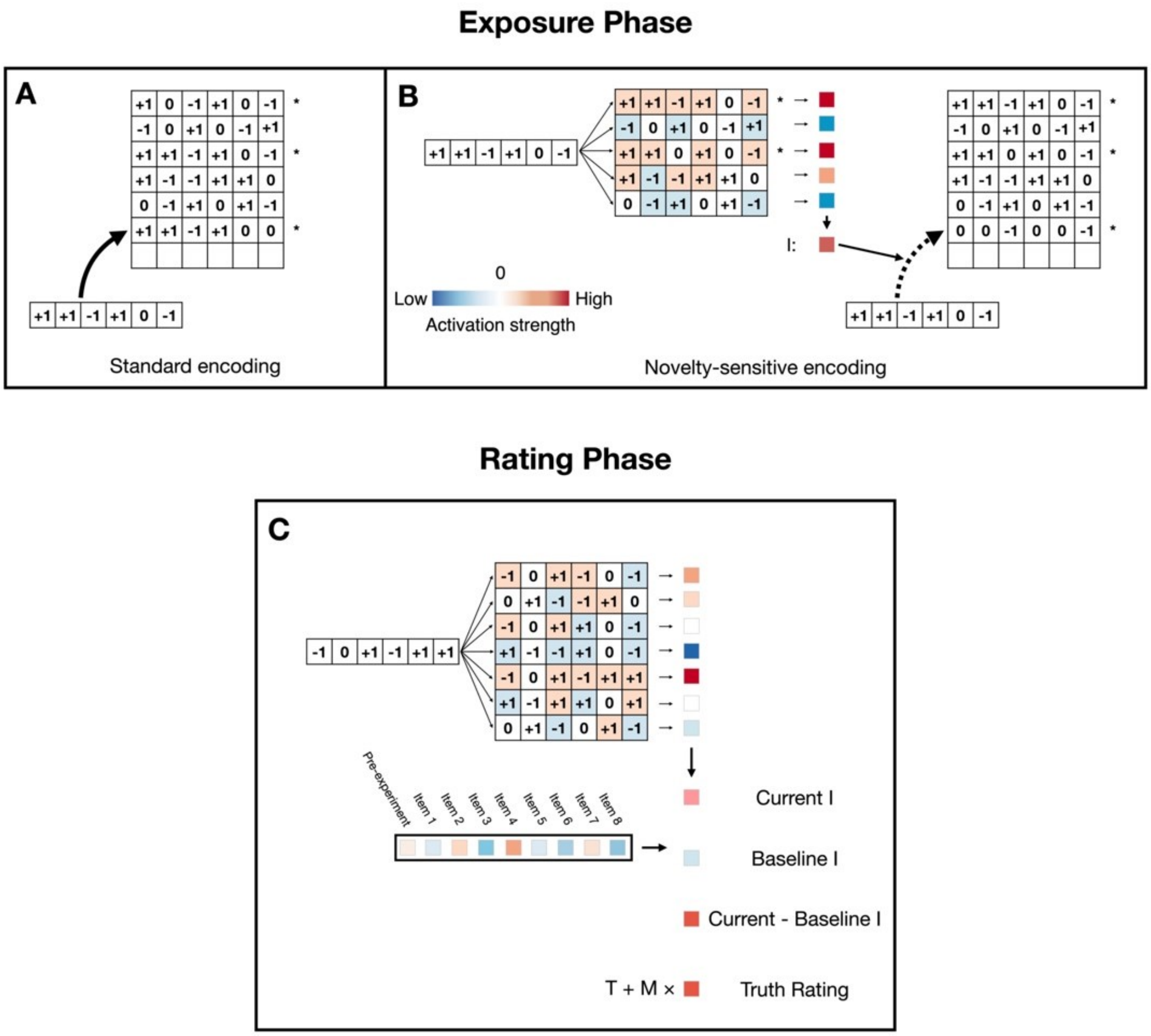

Figure 5. Schematic representation of the computational model. Panels A & B depict two ways of encoding items as traces into memory, here for an item that is being seen for the third time (asterisks represent all traces corresponding to encounters with the same item). In Panel A, each item is encoded with a standard learning rate (i.e., similar probability of features being transformed to 0), reflecting Model 1. In Panel B, encoding is novelty-sensitive, reflecting Model 2. The item is compared with all traces in memory, producing activation scores. Due to the existence of similar past items in memory, the intensity I (i.e., net activation) is positive, which decreases the learning rate. Panel C depicts how the model produces a truth rating, here for the ninth item in the truth rating phase. Intensity for a to-be-rated item (Current I) is compared with the average intensity of previous items (Baseline I). This difference in intensities is transformed

by a scaling parameter (M) then added to an average truth rating score (T) to produce the simulated truth rating.

Model 1

As mentioned earlier, Model 1 simply attempts to formalize the following two key theoretical assumptions: a) that repetition increases the familiarity and/or ease of processing associated with a statement (Unkelbach et al., 2019) and b) that the information from these cues is integrated into truth judgements in a relative fashion (i.e., greater than expected fluency increases truth judgements, and lower than expected fluency decreases truth judgements; Hansen et al., 2008; see Wänke & Hansen, 2015 for a review).

To implement the first assumption, this model is based on MINERVA 2 (Hintzman, 1984), an influential model of human episodic memory that has been used to model human performance in a variety of memory-related domains including recognition memory (Hintzman, 1988), schema abstraction (Hintzman, 1986) and the learning of word meanings (Jamieson et al., 2018).[4] Broadly, MINERVA 2 represents human memory as a database of representations of past experiences (here, the statements from the text messages). These representations, or traces, are created as partial copies of the experience (partial in the sense that no experience is perfectly encoded into memory). Familiarity arises when a probe (i.e., a current experience) is presented to this memory system: the probe is simultaneously compared with all of the traces in memory, returning an echo intensity. Echo intensity increases when there are many traces in memory that

[4] An adaptation of this model, MINERVA-DM (DM stands for decision-making), was also used to model a large variety of heuristics and biases involved in judgements of likelihood, including two early demonstrations of the illusory truth effect (Dougherty, Gettys, & Ogden, 1999 pp. 197-198). Unlike the present model, MINERVA-DM was primarily employed to compare repeated and new statements (rather than to predict the curve obtained with multiple repetitions) and was unable to represent more recent theoretical advances in the relationship between repetition-induced cues and judged truth (namely, that these cues are used in a relative fashion).

are similar to the probe, and when these traces are more similar to the probe. This echo intensity can be considered as a proxy for familiarity (Hintzman, 1986, 1988) or the experience of fluency.

Mathematically, the model represents each experience as a unique vector of length N, with each feature randomly taking on a value of +1, -1 (meaningful values) or 0 (an unknown or irrelevant feature). For each statement, a corresponding trace is formed by creating a partial copy of the vector and adding it to the bottom of an expanding matrix representing the human memory store. The fidelity of the partial copy is determined by the learning rate parameter L (set here to 0.6). Thus, the model records 60% of the existing features faithfully onto the memory trace and the remaining features are set to 0, indicating partial forgetting. Consistent with prior instantiations of MINERVA 2, this learning rate is set to a constant value. Given the success of MINERVA 2 in modelling the familiarity construct, and the critical role of familiarity in explanations of the illusory truth effect, this choice provides a justifiable baseline to test whether the key assumptions as presently stated in the literature can account for our findings.[5]

When presented with a probe, the echo intensity is calculated in several steps.[6] First, the model compares the probe item (another vector constructed in the same manner as a trace, here representing the to-be-rated statement) simultaneously with each trace stored in memory. Specifically, the model calculates a similarity score by tallying the number of identical non-zero features between the probe and a given trace and dividing this count by the total number of features that are non-zero in either the probe or the trace. In a second step, this similarity score for each trace is cubed to produce an activation score. Then, in a final step, these activation scores are summed across the memory store to produce the echo intensity. Note that this echo

[5] The precedent of a constant learning rate is also established in the MINERVA-DM model by Dougherty et al. (1999), who model the effects of three repetitions demonstrated by Hasher et al. (1977). This model is the the only prior demonstration of a computational model of the illusory truth known to the authors.
[6] See also Equations 1-3 in Hintzman (1984).

intensity must logically increase as a function of the number of similar traces to the probe stored in memory (reflecting the assumption that repeated exposure to a stimulus increases familiarity).

To reflect the second assumption (that repetition-induced cues are integrated into truth judgements in a relative fashion), we add a judgement component to the model described thus far. Specifically, truth ratings are modelled as a function of the difference between the echo intensity of the current to-be rated statement and a relative baseline. Mathematically, the relative baseline is the average of the previously rated items' echo intensities and an intensity value *B* which represents the baseline intensity at the start of the truth rating session. When the current statement's echo intensity is equal to the relative baseline, the model produces an average truth rating *T*. This rating increases or decreases by a scaling parameter (*M)* times the difference between the current echo intensity and the relative baseline. Note that this judgement component is a novel addition to existing MINERVA 2 models, the rest of the modeling framework is identical to MINERVA 2.

Model 2

Model 2 was identical to Model 1, with one simple modification to reflect the novelty-sensitive encoding assumption. Rather than the learning rate being set at a constant 0.6, the learning rate *L* was set to 0.6 minus the echo intensity of the current claim times a scaling parameter *S*. As a result, each successive repetition was encoded less faithfully, and thus contributed less to the final echo intensity during the rating phase.

Model comparison and discussion

For each model, we determined best-fitting parameters for *B*, *M*, and *T* (and *S*, for Model 2), given our data. Then, we used these parameters and the above models to simulate 500 subjects who first encoded the 10 key and 10 filler statements varying numbers of times and then provided truth ratings for the 10 key and eight new filler statements. Average truth ratings across simulated participants for each repetition schedule are shown in Figure 6. Statements were represented as vectors of length 100, with all features being randomly determined and all statement orders fully randomized. Full details of the modelling approach can be found at the project's OSF site: https://osf.io/re6dh.

The key question is whether a model reflecting only a few basic assumptions present in the illusory truth literature (Model 1) is sufficient to explain the present data, or if the addition of a novelty-sensitive encoding mechanism improves the performance of the model (Model 2). As shown in Figure 6, only Model 2 reproduced the pattern of results we obtained. A comparison of the Akaike's information criterion (AIC) for each model revealed substantial support for Model 2 (Model 1 AIC = 16,626.34, Model 2 AIC = 16,555.12). A comparison of Akaike weights (see e.g., Wagenmakers & Farrell, 2004) further suggests Model 2 is 2.93×10^{15} times more likely to be the best model than Model 1, providing strong evidence for the need for additional assumptions.

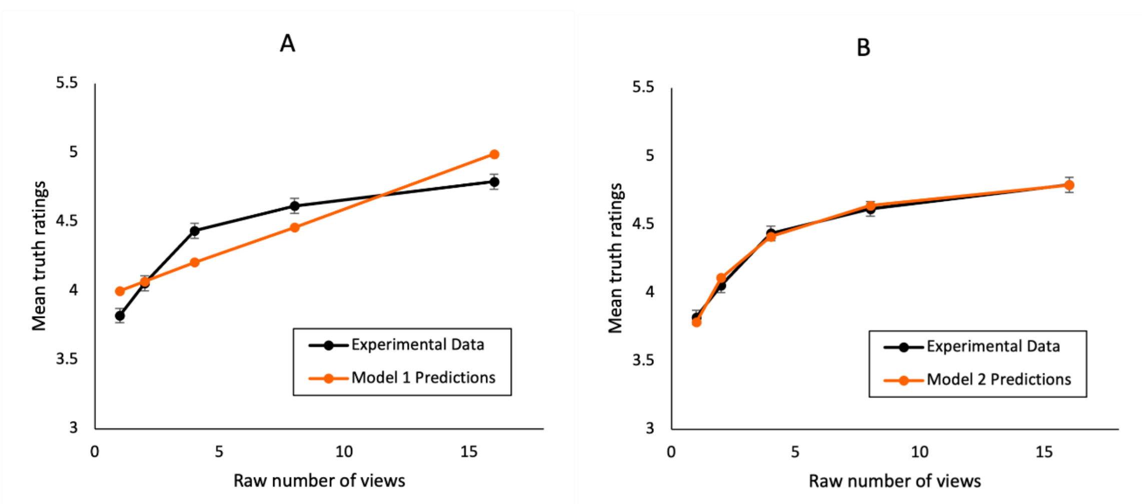

Figure 6. Mean truth ratings (collapsed across true and false statements) as a function of repetition in relation to (A) predictions from Model 1 and (B) predictions from Model 2. Error bars reflect standard error of the mean. Predictions are based on averages across 500 simulated

participants for each model using best-fitting parameter estimates for B, M, T (both models) and S (Model 2).

To be clear, the fact that the model with novelty-sensitive encoding fits the data well does not mean that is the correct cognitive mechanism. However, it is clear that the two theoretical features common in current theories are insufficient on their own to produce the current pattern of results. Some additional mechanism (possibly novelty-sensitive encoding) is required to produce the logarithmic curve. More broadly, the present work suggests that computational modelling is a useful tool for understanding the processes underlying the effects of repetition on truth judgements, and that experiments involving larger numbers of repetitions can shed light on aspects of these effects that may otherwise be unclear.

General Discussion

Overall, we find clear evidence that repetition increases belief in naturalistic settings. Repeated exposure to statements in participants' daily life increased their perceived truth. This increase was logarithmic – initial encounters produce the largest effects. However, additional repetitions continued to affect belief, even up to 16 exposures. In addition, we found no difference between spaced and massed repetitions. Repetition increased belief equally for statements that were repeated over consecutive days or spaced out across the two weeks.

However, it is important to note that spacing and timing were confounded within our design, so the massed presentations were also closer in time to the final test. It is possible that the spaced repetitions were more potent, but that the more recent exposure in the massed condition cancelled out those benefits. Future research should untangle the influence of both the spacing and recency of the repetitions. One possibility would be to include both massed schedules at the end of the presentation schedule (e.g., days 13-15) and at the beginning (e.g., days 1-3). While not a perfect solution, this method would ensure that the average delay between exposure and truth rating is equivalent across the spaced and massed groups. It is also true that even the "massed" presentations in our design only occurred once per day. Thus, it is perhaps more accurate to describe the two conditions as spacing (presentation every day) and greater spacing (presentation every few days). Regardless, we view the lack of a significant difference between the two conditions to be an interesting initial finding that deserves follow-up and, if replicated, has implications for theoretical explanations of the illusory truth effect.

The current findings have large implications for our daily life. Rather than only occurring in artificial laboratory situations, we find evidence that repetition increases belief in everyday activities. Exposure to true and false statements through text messages altered participants' beliefs. We know that many false rumors are continuously repeated in online spaces. For example, consider how many times you heard the "Obama is a Muslim" rumor during his presidency (see Kim & Kim, 2019 for correlational evidence that an increase in rumor circulation increased belief in this rumor). The current results provide experimental evidence that this type of real-world repetition can increase belief.

The results also have implications for theoretical explanations of the illusory truth effect. Our results fit with previous research finding a logarithmic relation between repetition and belief (DiFonzo et al., 2016; Hassan & Barber, 2021; Hawkins et al., 2001). Even with more naturalistic delays, we find that additional repetitions continue to affect belief, rather than quickly plateauing as in previous studies (Arkes et al., 1991). It is unclear exactly why our results differ from those of Arkes and colleagues (1991), however, one possibility is that in their study participants rated the truth of the statements after each exposure rather than only at the end of the study. Recent studies have shown that rating accuracy during exposure reduces effect of repetition on belief (Brashier et al., 2020; Calvillo & Smelter, 2020). When people have already thought about the accuracy of a statement it may reduce their reliance on peripheral cues such as fluency and familiarity, increasing the likelihood that they use their existing knowledge to judge the truth of the statement.

Finally, we introduce a simple computational model to test whether current theoretical assumptions (that repetition increases familiarity/fluency and that these cues are integrated relatively into ruth judgements) are sufficient to explain a logarithmic pattern of repetition effects. Current theories regarding the illusory truth effect have helped improve our understanding of how the mere process of seeing a statement can heighten a later truth judgement. However, as we demonstrate, current theoretical assumptions alone do not explain our pattern of results with multiple repetitions. We suggest that one additional assumption that can help bridge this gap is that successive repetitions are encoded less faithfully into memory (possibly due to decreased attention). This novelty-sensitive encoding mechanism is one that needs empirical testing. Regardless, it is one option that would produce the current pattern of results.

Conclusions

We found that repetition increased the perceived truth of trivia statements, even when the statements were encountered during participants' daily life outside of the laboratory. Importantly, the relation between repetition and perceived truth was logarithmic with early repetitions having a larger effect than later repetitions. Our simple computational models suggest that this logarithmic pattern can be explained by a novelty-sensitive encoding mechanism that better encodes new information (perhaps because people pay less attention to repeated information). Other mechanisms may actually be responsible, but it is clear that the existing explanation are not sufficient. Overall, our findings suggest that repetition affects our beliefs not only in experimental tasks, but also in real-world settings.

However, there are limits to the generalizability of our findings as well. Our repeated statements were bland trivia statements rather than the emotionally and politically charged types of misinformation that are often spread online. In addition, while participants were exposed to statements in their daily life, the exposure was more similar to encountering misinformation on WhatsApp or other messaging applications rather than in the context of a scrolling newsfeed. Now that we have established the viability of this research method, future studies should examine other types of materials and exposures.

Given that many interventions and experimental manipulations fail to generalize in more real-world situations (Blanco-Elorrieta & Pylkkänen, 2018; Hulleman & Cordray, 2009; Kingstone et al., 2003; Mitchell, 2012; Rendell & Craik, 2000), we think that it is essential that we attempt to move experimental psychology out of the lab and into real-life. The current study is an example of moving down that path. If we want psychological research to have relevance for real-world problems, we need to examine how our experimental effects play out in more naturalistic settings. The current results demonstrate that current worries about the repetition of false information in online spaces are valid (e.g., Kozyreva et al., 2020) and that repetition affects belief both inside and outside of the laboratory.

References

Arkes, H. R., Boehm, L. E., & Xu, G. (1991). Determinants of judged validity. *Journal of Experimental Social Psychology*, *27*(6), 576-605.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4 [sparse matrix methods; linear mixed models; penalized least squares; Cholesky decomposition]. *Journal of Statistical Software*, *67*(1), 48. https://doi.org/10.18637/jss.v067.i01

Blanco-Elorrieta, E., & Pylkkänen, L. (2018). Ecological Validity in Bilingualism Research and the Bilingual Advantage. *Trends in Cognitive Sciences*, *22*(12), 1117-1126. https://doi.org/https://doi.org/10.1016/j.tics.2018.10.001

Boehm, L. E. (1994). The validity effect: A search for mediating variables. *Personality and Social Psychology Bulletin*, *20*(3), 285-293.

Brashier, N. M., Eliseev, E. D., & Marsh, E. J. (2020). An initial accuracy focus prevents illusory truth. *Cognition*, *194*, 104054. https://doi.org/10.1016/j.cognition.2019.104054

Brown, A. S., & Nix, L. A. (1996). Turning lies into truths: Referential validation of falsehoods. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22*, 1088-1100.

Browning, K., & Alba, D. (2020). 'Tsunamis of Misinformation' Overwhelm Local Election Officials. *New York Times*. https://www.nytimes.com/2020/10/29/technology/misinformation-local-election-officials.html

Brumback, K., & Joffe-Block, J. (2020). Georgia poll worker in hiding after false claims online. *Associated Press*. https://apnews.com/article/fact-check-georgia-poll-worker-in-hiding-aa0f256ec21de96fd5a41da703c4b443

Calvillo, D. P., & Smelter, T. J. (2020). An initial accuracy focus reduces the effect of prior exposure on perceived accuracy of news headlines. *Cognitive Research: Principles and Implications*, *5*(1), 55. https://doi.org/10.1186/s41235-020-00257-y

Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The Truth About the Truth: A Meta-Analytic Review of the Truth Effect. *Personality and Social Psychology Review*, *14*(2), 238-257. https://doi.org/10.1177/1088868309352251

DiFonzo, N., Beckstead, J. W., Stupak, N., & Walders, K. (2016). Validity judgments of rumors heard multiple times: the shape of the truth effect. *Social Influence*, *11*(1), 22-39. https://doi.org/10.1080/15534510.2015.1137224

Dougherty, M. R., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, *106*(1).

Fazio, L. K. (2020). Repetition increases perceived truth even for known falsehoods. *Collabra: Psychology*, *6*, 38.

Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, *144*(5), 993-1002. https://doi.org/10.1037/xge0000098

Fazio, L. K., Rand, D. G., & Pennycook, G. (2019). Repetition increases perceived truth equally for plausible and implausible statements. *Psychonomic Bulletin & Review*, *26*(5), 1705–1710. https://doi.org/10.3758/s13423-019-01651-4

Gigerenzer, G. (1984). External validity of laboratory experiments: The frequency-validity relationship. *American Journal of Psychology*, *97*(2), 185-195.

Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493-498.

Hansen, J., Dechêne, A., & Wänke, M. (2008). Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology*, *44*, 687-691.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning & Verbal Behavior*, *16*, 107-112.

Hassan, A., & Barber, S. J. (2021). The effects of repetition frequency on the illusory truth effect. *Cognitive Research: Principles and Implications*, *6*(1), 1-12.

Hawkins, S. A., Hoch, S. J., & Meyers-Levy, J. (2001). Low-involvement learning: Repetition and coherence in familiarity and belief. *Journal of Consumer Psychology*, *11*(1), 1-11.

Henderson, E. L., Simons, D. J., & Barr, D. J. (2021). The Trajectory of Truth: A Longitudinal Study of the Illusory Truth Effect. *Journal of cognition*, *4*(1), 29-29. https://doi.org/10.5334/joc.161

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior research methods, instruments, & computers*, *16*, 96-101.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411-428.

Hintzman, D. L. (1988). Judgements of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*, 528 -551.

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, *2*(1), 88-110.

Jamieson, R. K., Avery, J. E., Johns, B. T., & Jones, M. N. (2018). An instance theory of semantic memory. *Computational Brain & Behavio*, *1*, 119-136. https://doi.org/10.1007/s42113-018-0008-2

Kim, J. W., & Kim, E. (2019). Identifying the Effect of Political Rumor Diffusion Using Variations in Survey Timing. *Quarterly Journal of Political Science*, *14*(3), 293-311.

Kingstone, A., Smilek, D., Ristic, J., Kelland Friesen, C., & Eastwood, J. D. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science*, *12*(5), 176-180.

Kozyreva, A., Lewandowsky, S., & Hertwig, R. (2020). Citizens versus the Internet: Confronting digital challenges with cognitive tools. *Psychological Science in the Public Interest*, *21*(3), 103-156. https://doi.org/10.1177/1529100620946707

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*(49), 433-442. https://doi.org/10.3758/s13428-016-0727-z

Mitchell, G. (2012). Revisiting truth or triviality: The external validity of research in the psychological laboratory. *Perspectives on Psychological Science*, *7*(2), 109-117.

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. . *Journal of Experimental Psychology: General*, *147*(12), 1865-1880. https://doi.org/10.1037/xge0000465

Pillai, R. F., L. . (2021). *Illusory Truth Texting Study - Trivia*. Retrieved from osf.io/re6dh

R Core Team. (2020). *R: A language and environment ofr statistical computing*. In https://www.R-project.org/.

Rendell, P. G., & Craik, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology*, *14*(7), S43-S62. https://doi.org/https://doi.org/10.1002/acp.770

Schwartz, M. (1982). Repetition and rated truth value of statements. *The American Journal of Psychology*, 393-407.

Unkelbach, C., & Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, *79*, 78-86.

Unkelbach, C., Koch, A., Silva, R. R., & Garcia-Marques, T. (2019). Truth by Repetition: Explanations and Implications. *Current Directions in Psychological Science*, *28*(3), 247-253. https://doi.org/10.1177/0963721419827854

Unkelbach, C., & Rom, S. C. (2017). A referential theory of the repetition-induced truth effect. *Cognition*, *160*, 110-126.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146-1151. https://doi.org/10.1126/science.aap9559

Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using akaike weights. *Psychonomic Bulletin & Review*, *11*, 192-196.

Wänke, M., & Hansen, J. (2015). Relative processing fluency. *Current Directions in Psychological Science*, *24*, 195-199. https://doi.org/10.1177/0963721414561766

Washington Post Fact Checker. (2020). Tracking all of President Trump's false or misleading claims. *The Washington Post*. https://www.washingtonpost.com/graphics/politics/trump-claims-database/